

# BUNDESREPUBLIK DEUTSCHLAND



## Bescheinigung

Herr Gerhard Weiglhofer in Weil, Kr Landsberg a Lech/Deutschland hat eine Patentanmeldung unter der Bezeichnung

"3D-Stereo-Echtzeit-Sensor-System"

am 24. November 1997 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig das Symbol G 06 T 15/00 der Internationalen Patentklassifikation erhalten.

München, den 4. Dezember 1998  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

Patentzeichen: 197 52 047.2

Ebert



# **TIEDTKE – BÜHLING – KINNE & PARTNER (GBR)**

Tiedtke-Bühling-Kinne, POB 20 19 18, D-80019 München

## **Patentanwälte Vertreter beim EPA\***

Dipl.-Ing. H. Tiedtke\*  
Dipl.-Chem. G. Bühling\*  
Dipl.-Ing. R. Kinne\*  
Dipl.-Ing. B. Pellmann\*  
Dipl.-Ing. K. Grams\*  
Dipl.-Biol. Dr. A. Link\*  
Dipl.-Ing. A. Vollnhals\*  
Dipl.-Ing. T. Leson\*  
Dipl.-Ing. H. Trösch\*  
Dipl.-Ing. Dr. G. Chivarov\*  
Dipl.-Ing. M. Grill\*  
Dipl.-Ing. A. Kühn

**Bavarlarling 4,  
D-80336 München**

24. November 1997

DE 20652

Gerhard Weiglhofer

Schwabhausen / Deutschland

**"3D-Stereo-Echtzeit-Sensor-System"**

Telefon: 0 89-54 46 90  
Telefax (G3): 0 89-53 26 11  
Telefax (G4): 0 89-53 29 09 50  
postoffice @ tbk-patent.com

Dresdner Bank (München) Kto. 3939 844 (BLZ 700 800 00)  
Deutsche Bank (München) Kto. 286 1060 (BLZ 700 700 10)  
Postbank (München) Kto. 670-43-804 (BLZ 700 100 80)  
Da-Ichi-Kangyo Bank (München) Kto. 51 042 (BLZ 700 207 00)  
Sanwa Bank (Düsseldorf) Kto. 500 047 (BLZ 301 307 00)

### Beschreibung:

Die vorliegende Erfindung betrifft eine Vorrichtung zur Bildverarbeitung, insbesondere eine Vorrichtung zur stereoskopischen Bildverarbeitung, sowie ein entsprechendes Verfahren.

Die optische Erfassung der dreidimensionalen Umgebung gewinnt seit geraumer Zeit zunehmend an Bedeutung. Insbesondere bei automatisierten Fertigungsabläufen und den dabei eingesetzten Robotern müssen neben zweidimensionalen Koordinaten auch Tiefeninformationen hinsichtlich eines (zu bearbeitenden) Objektes zur Verfügung stehen, um beispielsweise einen Roboter entsprechend steuern zu können.

Im Rahmen der "Ident Vision" Messe 1996 in Stuttgart wurde zum Beispiel ein Ganzkörperscanner vorgestellt, der mittels einer Vielzahl von Lasern und Kameras die räumliche Kontur eines Objekts erfassen kann. Jedoch ist dieses System für Echtzeit-Anwendungen ungeeignet.

Bislang bekannte Realisierungen zur Ermittlung von Tiefeninformationen kann man in verschiedene Gruppen unterteilen, die auf jeweils verschiedenen Verfahren beruhen.

Triangulationsverfahren verwenden eine oder mehrere Linien, die entweder von einem Laser oder einer Projektionseinheit auf das betreffende Objekt abgebildet werden. Mit einer oder mehreren Kameras wird das Objekt anschließend unter einem von der Projektionsrichtung geringfügig abweichenden, anderen Blickwinkel aufgenommen. Beruhend auf den vorbekannten geometrischen Verhältnissen erfolgt dann mittels trigonometrischer Berechnungen die Ermittlung der dreidimensionalen Struktur des Objektes.

Nachteilig bei diesen Verfahren ist jedoch, daß unter Umständen für jede projizierte Linie ein Bild auszuwerten ist, was äußerst zeitaufwendig ist, so daß ein derartiges System nicht echtzeittauglich ist.

Darüber hinaus ist es insbesondere nachteilig, daß das jeweilige Meßobjekt zu beleuchten ist. Dies stellt zwar unter Laborbedingungen kein Problem dar, wird aber in der Praxis fast unmöglich, wenn in Situationen wie im Straßenverkehr oder unter bei Teleskopen oder Mikroskopen vorliegenden erschweren realen Situationen Messungen durchzuführen sind. Ferner sind bei als Linienlichtprojektoren verwendeten Lasern spezielle Unfallverhütungsvorschriften zu beachten, damit bei derartigen Meßaufbauten keine Gefahr für das menschliche Auge besteht.

Bei einem speziellen Triangulationsverfahren wird ein Farbfächer auf das betreffende Objekt projiziert. Die von einer Kamera aufgenommenen Farbwerte können dann eindeutig einem Objektpunkt zugeordnet werden, wobei die Verschiebung einzelner Farbpunkte wiederum als Tiefeninformation auswertbar ist. Obwohl dieses Verfahren in Echtzeit arbeitet, bringt dieses Verfahren die einschränkende Forderung mit sich, daß das Objekt und der Hintergrund weiß sein müssen, so daß das Verfahren nicht universell einsetzbar ist.

Dagegen werten Laufzeitverfahren die Laufzeit von Licht, Mikrowellen oder eines Schallsignals als einem Meßsignal aus, um Rückschlüsse auf die räumliche Umgebung und die Tiefeninformation zu ziehen. Dabei wird jeweils eine Raumrichtung angepeilt und eine Abstandsmessung vorgenommen. Beispielsweise beruhen das Prinzip eines Radargerätes oder eines medizinischen Ultraschallgerätes auf derartigen Laufzeitverfahren. Laufzeitverfahren erfordern jedoch eine auf-

wendige Signalverarbeitung und sind zudem von den Eigenschaften des umgebenden Mediums abhängig, das die Ausbreitungseigenschaften des betreffenden Meßsignals beeinflusst.

5 Weiterhin kommen Schärfentiefeverfahren zum Einsatz, die jedoch nur dann angewendet werden können, wenn die Objektgröße und die Öffnung eines Objektivs in einem bestimmten Verhältnis zueinander stehen, wie beispielsweise in der Mikroskopie. Dabei wird durch das optische System nur der  
10 schmale Bereich der Fokalebene scharf auf einen Bildsensor abgebildet. Die Bereiche davor und dahinter sind mehr oder weniger unfokussiert. Die in der Fokalebene enthaltene Bildinformation kann mit herkömmlichen Bildverarbeitungsverfahren ausgewertet werden. Eine entsprechende Auswertung  
15 liefert dann eine Höhenlinie des Objekts. Durch mehrfaches Verschieben der Fokalebene und anschließender Auswertung kann dann eine Höhenlinienkarte des Objekts erstellt werden, was jedoch entsprechend aufwendig ist.

20 Ein weiterer Ansatz, der erst seit wenigen Jahren verfolgt wird, beruht in der Auswertung stereoskopischer Bilder zum Gewinnen von Tiefeninformationen, ähnlich dem räumlichen Sehen beim Menschen, dessen Nachbildung durch derartige Verfahren angestrebt ist.

25 Das Phänomen des räumlichen Sehens, das nachfolgend kurz in Bezug auf den Menschen allgemein erläutert ist, beruht auf der Wahrnehmung der sichtbaren Umwelt mittels der beiden menschlichen Augen, wobei aufgrund des Augenabstandes die  
30 beiden wahrgenommenen Bilder (rechtes Auge bzw. Bild, linkes Auge bzw. Bild) unter leicht unterschiedlichen Blickwinkeln wahrgenommen werden. Da die Sehachsen der Augen etwas konvergieren, treffen sich ihre Sehachsen an einem Punkt der betrachteten Umwelt, wobei von einer Fixierung  
35 dieses Punktes durch die Augen gesprochen wird. Das Bild

dieses Punktes fällt jeweils auf eine Sehgrube der Netzhaut. Jeder benachbarte Punkt im Gesichtsfeld wird dann auf eine Stelle der Netzhaut projiziert, die etwas vom Zentrum des schärfsten Sehens entfernt liegt. Im allgemeinen ist

5 dieser Abstand in beiden Augen unterschiedlich, wobei der Abstand als Funktion der räumlichen Tiefe des betrachteten Punktes im Gesichtsfeld relativ zum Fixationspunkt variiert. Diese Abstandsabweichungen werden als binokulare Disparitäten bezeichnet, die nachfolgend kurz als

10 "Disparitäten" bezeichnet sind.



Beim Stereosehen besteht das eigentliche Problem, neben der Auswertung der jeweils gegebenen trigonometrischen Gegebenheiten darin herauszufinden, welcher Bildpunkt der realen Welt bzw. des gesehenen Objekts, der im ersten der stereoskopischen Bilder auftritt, welchem Bildpunkt im zweiten der stereoskopischen Bilder entspricht. Anders ausgedrückt, gilt es in Bezug auf das menschliche Auge herauszufinden, welcher Bildpunkt auf der linken Netzhaut welchem Bildpunkt

15 auf der rechten Netzhaut entspricht.

20



Für das räumliche Sehen bzw. die 3D-Stereowahrnehmung ist es dann erforderlich, aus den Disparitäten auf die dreidimensionale Struktur des wahrgenommenen Bildes zurück zu schließen. Anders ausgedrückt, aus ermittelten Disparitäten kann auf die in dem wahrgenommenen Bild enthaltenen Tiefeninformationen geschlossen werden.

25

Bislang verfügbare Stereoverfahren sind jedoch auf leistungsfähige Arbeitsplatzrechner (PC's bzw. sogenannte "Workstations") angewiesen, wobei zur Ermittlung der gewünschten Tiefeninformationen aus einem einzigen Stereobildpaar, beispielsweise unter Verwendung eines handelsüblichen Arbeitsplatzrechners (mit einer Taktfrequenz von et-

30

wa 90 MHz), eine Zeit im Bereich von etwa 3 bis 5 Minuten erforderlich ist.

Versucht man jedoch, derartige Verfahren in spezielle Rechenwerke auszulagern, übersteigt der Entwicklungs- und Hardwareaufwand sehr schnell wirtschaftlich vertretbare Grenzen. Die Markteinführung derartiger Systeme scheiterte bislang also entweder an einer zu geringen Verarbeitungsgeschwindigkeit oder einem zu hohen Systempreis.

Stereoverfahren sind bestrebt, einander entsprechende Bildpunkte zwischen einem ersten (linken) und zweiten (rechten) Stereobild zu finden, wobei dann aus der relativen Verschiebung solcher Bildpunkte bei bekannter Kameraposition die Entfernung der Punkte berechnet werden kann.

Unter den vorstehend angesprochenen herkömmlichen rechnerimplementierbaren Stereoverfahren kann man im wesentlichen unterscheiden zwischen sogenannten merkmalsbasierten Stereoverfahren, intensitätsbasierten Stereoverfahren und phasenbasierten Stereoverfahren.

Bei merkmalsbasierten Stereoverfahren werden vor der Korrespondenz-Suche bestimmte Merkmale (z.B. Kanten oder Intensitätsmaxima) aus den Bilddaten extrahiert, die für stabiler als die rohen Bildintensitäten gehalten werden. Derartige Verfahren sind stabil, falls nur wenige stabile Merkmale (etwa Objektkanten) extrahiert werden. Sie sind auch schnell, da Bilder nur an den Orten weiterverarbeitet werden müssen, an denen Merkmale erfaßt wurden. Dies führt zu einer merklichen Datenreduktion.

Jedoch können dadurch Disparitäten nur an den Stellen berechnet werden, an denen die gewählten Merkmale erfaßt wurden. An allen weiteren Stellen des Bildes muß interpoliert

werden, was zusätzlich zeitaufwendige Rechenvorgänge notwendig macht. Je dichter die Merkmale liegen, desto schwieriger wird es, diese einander zuzuordnen. Die endgültige Zuordnung kann nur anhand komplizierter einzusetzender Verfahren getroffen werden, was zu einer verminderten Verarbeitungsgeschwindigkeit führt.

Intensitätsbasierte Stereoverfahren arbeiten direkt mit den seitens der Stereokameras bzw. der Bildaufnehmer gelieferten Bildhelligkeiten. Daher sind zur Zuordnung einander entsprechender Helligkeiten sehr stabile Algorithmen einzusetzen, die entsprechend viel Rechenzeit erfordern. Unter den intensitätsbasierten Verfahren kann weiter unterschieden werden zwischen korrelationsbasierten Verfahren und Verfahren mittels dynamischer Programmierung.

Erstere sind bestrebt, einen kleinen Bildbereich des linken Bildes in dem rechten Bild wiederzufinden. Dazu werden üblicherweise Verfahren zur Maximierung der Korrelation oder Minimierung der Differenz eingesetzt. Falls die Bildausschnitte groß genug gewählt sind, erzielt man mit diesen Verfahren stabile und dichte Disparitätskarten, wobei eine Echtzeitverarbeitung unter Einsatz spezieller aufwendiger Hardware denkbar ist. Die rohen Disparitätskarten enthalten typischerweise jedoch nur Verschiebungen im Bereich der Genauigkeit eines Bildelements (Pixel), so daß die Auflösung begrenzt ist. Nur durch Einsatz von Interpolationsverfahren oder aber iterativen Verfahren kann die Genauigkeit der Disparitäten jedoch auf subpixelgenaue Disparitäten verbessert werden.

Bei Verfahren mittels dynamischer Programmierung wird versucht, lokal die Differenz zu minimieren, wobei diese lokale Minimierung dann anhand geeigneter Verfahren auf das gesamte Bild ausgedehnt wird. Typischerweise berechnen auch



diese Verfahren Disparitäten nur auf ein Bildelement genau, was für zahlreiche Anwendungen jedoch zu ungenau ist.

Zusammenfassend ist festzustellen, daß intensitätsbasierte Verfahren zwar eine dichte Disparitätskarte (einen Disparitätswert für fast jeden Bildpunkt) liefern, die Minimierung des Fehlermaßes wie bei dem Verfahren der Minimierung der Differenz jedoch zeitaufwendig ist und daher nicht für Echtzeitanwendungen geeignet ist. Die Reduzierung der erforderlichen Rechenzeit kann durch Einsatz hierarchischer Verfahren reduziert werden, dies jedoch nur durch Einsatz komplizierter und nicht parallelisierbarer algorithmischer Strukturen, für die eine Hardwarelösung zumindest äußerst aufwendig und somit kostenintensiv wird.

Desweiteren sind phasenbasierte Stereoverfahren Gegenstand aktueller Untersuchungen. Diese Verfahren extrahieren aus den rohen Bilddaten eine lokale Fourier-Phase (meist durch Filterung mittels Gabor-Funktionen), und versuchen dann, diese Fourier-Phasenbilder (für das rechte bzw. linke Bild) in Übereinstimmung zu bringen. Die Extraktion der lokalen Fourier-Phase entspricht dabei einem lokalen Kontrast-Ausgleich, wodurch viele durch Helligkeitsschwankungen verursachte Fehler anderer Stereoverfahren reduziert werden. Insbesondere entspricht die Berechnung der lokalen Fourier-Phase einer Merkmalsextraktion, wobei das Merkmal (die Phase) kontinuierlich verteilt ist, also an jedem Bildpunkt bzw. Bildelement vorhanden ist. Diese phasenbasierten Stereoverfahren arbeiten subpixel-genau und sind auch für Echtzeitanwendungen implementierbar. Jedoch müssen zur Erzeugung dichter Disparitätskarten hierarchische Verfahren eingesetzt werden, die zur Fehlervermeidung meist spezielle Ausnahmesituationen abfangen müssen. Dies macht die verwendeten Algorithmen kompliziert und eine Hardwarerealisierung aufwendig und kostenintensiv.

Es ist folglich Aufgabe der vorliegenden Erfindung, eine Vorrichtung und ein Verfahren zur Bildverarbeitung zu schaffen, mit der bzw. dem stereoskopische Bilder in Echtzeit verarbeitet werden können, und die bzw. das mit einfachen Mitteln und geringem Aufwand zu realisieren ist.

Erfindungsgemäß wird diese Aufgabe gelöst durch eine Vorrichtung zur Bildverarbeitung, mit Verschiebe-Einrichtungen zur gegenseitigen Verschiebung von diesen jeweils zugeführten Bilddaten eines ersten Bildes und eines zu diesem unter einem anderen Aufnahmewinkel aufgenommenen zweiten Bildes; einer Disparitäts-Detektionseinrichtung zur Erfassung von räumlichen Tiefeninformationen wiedergebenden Disparitätswerten für durch die Verschiebe-Einrichtungen gegeneinander verschobenen Bilddaten des ersten Bildes und des zweiten Bildes; und einer Kohärenz-Detektionseinrichtung zur Ermittlung des tatsächlichen Disparitätswertes unter den Disparitätswerten dadurch, daß in einem vorbestimmten Wertebereich um einen betreffenden der Disparitätswerte eine größtmögliche Anzahl der übrigen Disparitätswerte festgestellt wird.

Ferner wird diese Aufgabe gelöst durch ein Verfahren zur Bildverarbeitung, mit den Schritten: gegenseitiges Verschieben von zugeführten Bilddaten eines ersten Bildes und eines zu diesem unter einem unterschiedlichen Aufnahmewinkel aufgenommenen zweiten Bildes mittels jeweiliger Verschiebe-Einrichtungen; Erfassen von räumliche Tiefeninformationen wiedergebenden Disparitätswerten für die gegeneinander verschobenen Bilddaten des ersten Bildes und des zweiten Bildes mittels einer Disparitäts-Detektionseinrichtung; und Ermitteln des tatsächlichen Disparitätswertes unter den Disparitätswerten derart, daß in einem vorbestimmten Wertebereich um einen betreffenden der Dispari-

tätswerte eine größtmögliche Anzahl der übrigen Disparitätswerte festgestellt wird.

5 Durch diesen erfindungsgemäßen Aufbau des Systems bzw. diese Verarbeitungsweise des erfindungsgemäßen Verfahrens können die zugeführten Bildsignale mit der Geschwindigkeit verarbeitet werden, wie sie von Bildaufnehmern geliefert werden. Somit ist die Bilddatenverarbeitung zur Gewinnung von Tiefeninformationen mit geringem Aufwand und in Echtzeit möglich.

10 Diese hohe Verarbeitungsgeschwindigkeit (Echtzeit) des Systems ist ein wesentlicher Vorteil gegenüber dem Stand der Technik. Mit einer Bildrate von beispielsweise 25 Bildern pro Sekunde, die dem System seitens der Bildaufnehmer zugeführt werden und von diesem in Echtzeit verarbeitet werden, ist das System bzw. das mit diesem implementierte Verfahren den aus dem Stand der Technik bekannten Anordnungen bzw. Verfahren zumindest um einen Geschwindigkeitsfaktor von 100 überlegen.

25 Weiterhin ist das zu messende Volumen des Meßobjekts gegenüber bekannten Systemen nicht mehr begrenzt. Insbesondere ist der Meßbereich des erfindungsgemäßen Systems nur durch die Begrenzung der verwendeten abbildenden Optik bestimmt, und alle optischen Vorsatzgeräte wie Mikroskope, Tele- oder Weitwinkelobjektive können in Verbindung mit den Bildaufnehmern eingesetzt werden, sofern sie an diese angepaßt sind. Insbesondere ist es mit dem erfindungsgemäßen System möglich, ein Meßobjekt ohne Markierungshilfen auszuwerten, die bei Systemen gemäß dem Stand der Technik häufig erforderlich sind.

30 Durch die Möglichkeit ein Meßobjekt ohne Markierungshilfen auszuwerten entfällt weiterhin die Notwendigkeit eine dafür

notwendige Projektionseinheit vorzusehen. Folglich wird das erfindungsgemäße System kompakter und leichter, wobei sich die Handhabung erleichtert und sich vielseitigere Einsatzmöglichkeiten ergeben.

5

Zudem entfällt bei dem erfindungsgemäßen System auch eine Vorbehandlung des Meßobjekts selbst, beispielsweise ein Abdecken des Meßobjekts mit hellem Puder zur Beseitigung von Spiegelungen und Glanzlichtern auf der Objektoberfläche vor der Messung und Beseitigung des Puders nach erfolgter Messung. Durch die Einsparung derartiger Arbeitsgänge arbeitet das erfindungsgemäße System / Verfahren zeit- und kostengünstiger.

10

Zudem ist es bei dem erfindungsgemäßen System vorteilhafterweise möglich, beliebige Bildsignalquellen anzuschließen. Das heißt, das erfindungsgemäße System bzw. Verfahren ist nicht auf den Bereich sichtbaren Lichts beschränkt, sondern die Auswertung kann beispielsweise auch für Bildsignale im Infrarotbereich erfolgen.

20

Vorteilhafte Weiterbildungen der Erfindung sind in den Unteransprüchen angegeben.

25

Die vorliegende Erfindung ist nachstehend anhand von Ausführungsbeispielen mit Bezug auf die beigefügte Zeichnung näher beschrieben. Dabei zeigen:

30

Fig. 1 ein allgemeines Blockschaltbild des erfindungsgemäßen Systems;

35

Fig. 2 ein Blockschaltbild, das Bildsignal-Verschiebe-  
einrichtungen und Disparitätsbestimmungseinrichtungen betreffende Einzelheiten zeigt;

Fig. 3A ein Funktionsschaltbild eines einzelnen gradientenbasierten bzw. differentiellen Disparitätselements;

Fig. 3B ein Blockschaltbild zur Hardwarerealisierung des gemäß Fig. 3A funktional beschriebenen einzelnen Disparitätselements.

Fig. 4A ein Blockschaltbild eines Beispiels eines Kohärenzdetektors;

Fig. 4B eine schematische Darstellung einer Ausführungsvariante eines Kohärenzdetektors;

Fig. 5 ein Diagramm, das für alle Bildelemente einer Bildzeile die von verschiedenen Disparitätselementen ermittelten Disparitäten darstellt, aus denen dann mittels eines Kohärenzdetektors in Echtzeit die tatsächliche bzw. wahre Disparität im Verlauf der Bildzeile bestimmt wird.

Fig. 6 ein Beispiel zweier als Bildsignalstrom zugeführter Bilder eines ersten und zweiten (bzw. linken und rechten) Bildaufnehmers, sowie eine Darstellung der mit dem erfindungsgemäßen System ermittelten Tiefeninformation in Form einer sogenannten Disparitätskarte.

Fig. 1 stellt das erfindungsgemäße stereoskopische Sensorsystem zur Erfassung dreidimensionaler Objekte in Echtzeit schematisch als Blockschaltbild dar.

Das System weist einen Sensorkopf auf, der zwei Bildaufnehmer umfaßt, einen Bildaufnehmer CCDL für das (erste) linke Bild und einen Bildaufnehmer CCDR für das (zweite) rechte Bild. Bei den Bildaufnehmern kann es sich beispielsweise um ein Kamerapaar mit CCD-Bildaufnehmern handeln, die in einem vorbestimmten bekannten Abstand  $d$  voneinander angeordnet sind und daher leicht unterschiedliche Bilder des selben Objektes aufnehmen (vgl. Fig. 6). Insbesondere können die

Bildaufnehmer Bewegtbilder aufnehmen und diese dem System zur nachfolgenden Echtzeitverarbeitung zuführen. Mittels einer Bildaufnehmer-Synchronisationseinrichtung SYNC werden die Aufnahmen der Bildaufnehmereinrichtungen CCDL, CC DR

5 zeitlich synchronisiert, so daß das linke und das rechte Bild synchron zueinander aufgenommen werden. Je nach Entfernung des Objekts von den Bildaufnehmern bzw. Kameras weisen die aufgenommenen Bildpunkte leichte - als Disparitäten bezeichnete - Verschiebungen auf. Mittels einer nachfolgend als Disparitätsdetektor DD bezeichneten Einrichtung und einer als Disparitäts-Kohärenzdetektor KD bezeichneten Einrichtung (die beide nachfolgend noch ausführlich beschrieben sind) können die tatsächlichen Disparitäten bzw. Verschiebungen ermittelt werden und dann durch trigonometrische Berechnungen unter Berücksichtigung des bekannten Abstandes  $d$  zwischen den Bildaufnehmern CCDL, CC DR die räumliche Entfernung des jeweiligen Objekts berechnet werden, somit die gesuchten Tiefeninformationen erhalten werden. Die so aus zwei Stereo-Bewegtbildern berechnete Tiefeninformation kann quasi verzögerungsfrei in Echtzeit als Videosignal ausgegeben werden und als dritte Dimension (z. B. drittes Bild) als Grauwertbild oder Farbbild dargestellt werden. Dabei ist die Echtzeitverarbeitung des Systems insbesondere durch die serielle Verarbeitung der durch die

10  
15  
20  
25

Mit der Bildaufnehmer-Synchronisationseinrichtung SYNC verbunden ist ein Taktgenerator CLK, der Taktsignale zur Steuerung und Synchronisation weiterer Komponenten des Systems erzeugt.

30

Die aufgenommenen Bildsignale werden von den Bildaufnehmern CCDL, CC DR diesen jeweils zugeordneten Vorverarbeitungseinrichtungen VVL, VVR zugeführt. In diesen Vorverarbeitungseinrichtungen VVL, VVR erfolgt ein Helligkeitsausgleich der

35

beiden Stereobilder vor der nachfolgenden Verarbeitung. Diese Helligkeitsvorverarbeitung ist vorteilhaft, da aufgrund leicht unterschiedlicher Bildaufnahmewinkel der Bildaufnehmer CCDL, CCDDR sogenannte Glanzlichter und Reflexionen an (spiegelnden) Oberflächen der aufgenommenen Objekte auftreten können, die bei der nachfolgenden Verarbeitung der Bilddaten zur Erfassung der Disparitäten zu Fehlern führen können, so daß die Disparitätswerte instabil werden könnten.

10

Für den Fall der Verarbeitung von Farbbildern kann im Rahmen der Vorverarbeitung durch die Vorverarbeitungseinrichtungen eine Zerlegung der jeweiligen (ersten und zweiten) Farbbilddaten in jeweilige Farbbildkomponenten (z. B. in die Primärfarben Rot, Grün, Blau (RGB) oder aber in die Luminanz- und die beiden Farbdifferenzkomponenten) erfolgen. Zudem ist es auch möglich, die ersten und zweiten Bilddaten unterschiedlichen Vorverarbeitungen zu unterziehen.

20

Nach erfolgter Vorverarbeitung der Bilddaten seitens der Vorverarbeitungseinrichtungen VVL, VVR werden die Bilddaten des ersten und zweiten (bzw. linken und rechten) Bildes jeweils Einrichtungen VSEL, VSER zur gegenseitigen Verschiebung der linken und rechten Bilddaten in Bezug aufeinander zugeführt. Genauer heißt das, daß jeweils zwei Verschiebeeinrichtungen VSEL, VSER eine Verschiebeeinheit bilden, wobei dabei je ein Verschiebestrom von Bilddaten nach rechts und nach links pro Verschiebeeinheit vorliegt. Eine Verschiebeeinheit ist dabei je nach Anwendungsfall für jede Art von Bilddaten vorzusehen. Das heißt, im Fall einer Vorverarbeitung durch Zerlegung in die Primärfarben R, G und B können drei Verschiebeeinheiten vorgesehen sein, und zwar je eine für jede Primärfarbe.

35


Die Disparitäten derart zueinander verschobener Bilddaten werden dann für jeweilige Bilddatenpaare durch den Disparitätsdetektor DD ermittelt und ein jeweiliger Satz erhaltener Disparitätswerte für die Bilddatenpaare wird durch den  
5 Disparitäts-Kohärenzdetektor KD ausgewertet. Diese Kohärenzdetektion bezüglich der erhaltenen Disparitäten beruht auf der Auswahl bzw. Ermittlung desjenigen der erhaltenen Disparitätswerte, der annähernd identisch mit zumindest einem weiteren der erhaltenen Disparitätswerte ist. Anders  
10 ausgedrückt wird der Disparitätswert als der eigentliche bzw. wahre Disparitätswert ermittelt, der in etwa identisch zu zumindest einem weiteren Disparitätswert (d.h. in einem vorbestimmten Intervall um diesen liegt) oder identisch mit zumindest einem weiteren Disparitätswert ist. Fig. 5 veranschaulicht dieses Prinzip anhand von Daten von beispielsweise sieben Disparitätselementen eines Disparitätsdetektors.

Die seitens des Disparitäts-Kohärenzdetektors KD derart ermittelten eigentlichen Disparitätswerte werden dann von  
20 diesem an eine Einrichtung NB zur Bilddatennachbearbeitung ausgegeben. Diese Einrichtung NB bereitet die Disparitätswerte so auf, daß diese über einen Ausgangsanschluß DV einer wahlweise anzuschließenden Ausgabevorrichtung wie zum Beispiel einer (nicht dargestellten) Anzeigeeinrichtung zugeführt werden können und von dieser als Grauwert- oder  
25 Farbbild wiedergegeben werden. Zur besseren Darstellbarkeit der Disparitätswerte erfolgt im Rahmen der Bilddatennachbearbeitung auch eine Beseitigung auftretender Randeffekte  
30 sowie eine virtuelle Begrenzung.


In der Figur selbst nicht dargestellt ist eine externe Bedieneinheit zur Konfiguration des Systems, die durch einen Arbeitsplatzrechner mit geeigneter Schnittstelle zu dem System implementiert sein kann.  
35



Fig. 2 der Zeichnung zeigt nunmehr Einzelheiten der Bilddaten-Verschiebeeinrichtung VSEL für Bilddaten des (ersten) linken Bildes, der Bilddaten-Verschiebeeinrichtung VSER für Bilddaten des (zweiten) rechten Bildes, die zusammen eine Verschiebeeinheit bilden, und des Disparitätsdetektors DD als Disparitätsbestimmungseinrichtung. Dabei sind die Systemkomponenten gemäß Fig. 2 nicht auf eine bestimmte Implementierung in analoger oder digitaler Technik festgelegt und ihre Beschreibung erfolgt zunächst in allgemeiner Form, unabhängig von der tatsächlichen Realisierungsform.



Die von den jeweiligen Vorverarbeitungseinrichtungen VVL, VVR ausgegebenen Bilddaten werden den Bilddaten-Verschiebeeinrichtungen VSEL, VSER zugeführt. Im gezeigten Fall ist dabei zur Vereinfachung der Darstellung nur eine Verschiebeeinheit für eine Art von Bilddaten, beispielsweise für eine der Primärfarben R, G, B dargestellt.



Dabei werden die jeweiligen Bilddaten diesen Verschiebeeinrichtungen der Verschiebeeinheit jeweils seriell und synchron zueinander zugeführt. Jede der Verschiebeeinrichtungen VSEL, VSER besteht aus einer Vielzahl von Verschiebeelementen  $\Delta X_{L1}$  bis  $\Delta X_{Ln}$  bzw.  $\Delta X_{R1}$  bis  $\Delta X_{Rn}$ . Diese Elemente bewirken jeweils eine Verschiebung der zugeführten Bildelemente um eine kleine Strecke in x-Richtung, d.h. in Richtung der Hauptabtastrichtung bei zeilenweise abtastenden Bildaufnehmern. (Fig. 2 zeigt einen Fall für  $n=7$ .)

Beginnend mit dem jeweils ersten Bildelement der seriellen Bilddatenströme des linken bzw. rechten Bildes (z.B. Bildelement der oberen linken Ecke des jeweiligen Bildes) werden die Bildelemente (Pixel) gegeneinander verschoben. Die Auswertung der Bildelemente hinsichtlich der vorhandenen Disparität erfolgt dann jeweils paarweise. Genauer erfolgt

eine paarweise Auswertung hinsichtlich der Disparität beruhend auf den verschobenen Bilddaten für Paare  $(\Delta X_{L1}, \Delta X_{Rn}), (\Delta X_{L2}, \Delta X_{Rn-1}), \dots, (\Delta X_{Ln-1}, \Delta X_{R2}), (\Delta X_{Ln}, \Delta X_{R1})$ . Der Betrag  $\Delta X$  der Verschiebung kann auf einen Wert eingestellt werden, der der Länge eines Bildelements entspricht. Falls eine höhere Genauigkeit (Subpixel-Genauigkeit) gewünscht ist, kann diese Verschiebung jedoch auch auf einen Wert eingestellt werden, der kleiner als die Länge eines Bildelements ist, beispielsweise auf die halbe Länge eines Bildelements (Pixels).

Im Fall von als Analogschaltung realisierten Verschiebeeinrichtungen VSEL, VSER bestehen diese jeweils aus einer Kette bzw. Kaskade von Laufzeitgliedern (Verzögerungsgliedern) mit entsprechenden Signalabgriffen. Der Verschiebungsbetrag in x-Richtung entspricht dann der Laufzeit und kann bild-elementweise oder in Bruchteilen von Bildelementen gewählt sein.

Im Fall von als Digitalschaltung realisierten Verschiebeeinrichtungen VSEL, VSER besteht die jeweilige Einheit aus in Reihe geschalteten Schieberegistern, denen ein entsprechendes Taktsignal von dem Taktgenerator CLK (ggfs. über einen zwischengeschalteten (nicht dargestellten) Frequenzteiler) zuzuführen ist. Dann ist es jedoch zuvor erforderlich, die analogen Bilddaten mittels eines Analog/Digital-Wandlers in digitale Bilddaten umzuwandeln. Ein derartiger Analog/Digital-Wandler kann vor oder nach der jeweiligen Vorverarbeitungseinrichtung VVL, VVR im Signalverlauf vorgesehen sein, je nachdem, ob die Vorverarbeitung noch analog oder schon digital erfolgt.

Im Fall der digitalen Schieberegister erfolgt die Ansteuerung der Verschiebeelemente  $\Delta X_{Li}, \Delta X_{Ri}$  der jeweiligen Ver-

schiebeeinrichtung zur bildelementweisen Verschiebung jeweils entsprechend dem Takt, mit dem die digitalen Bilddaten zugeführt werden. Ist Subpixel-Genauigkeit erforderlich, so sind die bereits digitalisierten Daten einem speziellen Filter zuzuführen. Zum Beispiel wird eine einem halben Bildelement entsprechende Verschiebung erreicht, indem man zwei aufeinanderfolgende Bildelemente mittels eines entsprechenden Filters einer Mittelwertbildung unterzieht und den erhaltenen Wert als Signalwert des Bildelements an der um ein halbes Bildelement verschobenen Position annimmt. Eine derartige Filterung entspricht dann im wesentlichen einer Ditherung benachbarter Bildelemente.

Die vorstehend genannten Bilddatenpaare  $(\Delta X_{L1}, \Delta X_{Rn})$ ,  $(\Delta X_{L2}, \Delta X_{Rn-1})$ , ...,  $(\Delta X_{Ln-1}, \Delta X_{R2})$ ,  $(\Delta X_{Ln}, \Delta X_{R1})$  werden dann jeweils einem Element  $EST_1, \dots, EST_n$  des Disparitätsdetektors DD zur Auswertung zugeführt, um die entsprechende Disparität für das jeweilige Bilddatenpaar zu ermitteln. Es ist zu beachten, daß aufgrund der vereinfachten Darstellung gemäß Fig. 2 nur ein Disparitätsdetektor DD dargestellt ist. Sind jedoch mehrere Verschiebeeinheiten (z.B. jeweils für R-, G-, B- Bilddaten vorhanden), so ist jeder der Verschiebeeinheiten ein entsprechender Disparitätsdetektor zuzuordnen.

Für jeweilige Disparitäts-Elemente  $EST_i$  eines Disparitätsdetektors DD sind verschiedene Arten denkbar. Beispielsweise kann die Auswertung gradientenbasiert, tensorbasiert, energiebasiert oder aber gemäß dem Reichard'schen Modell erfolgen. Diese Auswerteprinzipien sind in der Fachliteratur hinreichend beschrieben und es wird hier nicht mehr ausführlich darauf eingegangen. Hinsichtlich der erzielbaren Genauigkeit und eines einfachen Schaltungsaufbaus sind gradientenbasierte Disparitäts-Elemente den tensor- bzw. energiebasierten vorzuziehen.

Die von den Disparitäts-Elementen  $EST_i$  ( $1 \leq i \leq n$ ) ausgegebenen Disparitätswerte werden nachfolgend dem Kohärenzdetektor KD zugeführt, der aus den zugeführten Werten den eigentlichen Disparitätswert bestimmt. Insbesondere im Fall mehrerer Disparitätsdetektoren DD werden alle Ausgänge der jeweiligen Disparitätsdetektoren einem nachfolgenden Kohärenzdetektor KD zugeführt.

Fig. 3A zeigt nunmehr ausführlich ein Funktionsschaltbild eines Disparitäts-Elements  $EST_i$ , bei dem die Auswertung auf dem gradientenbasierten Ansatz beruht.

Dabei werden dem jeweiligen Disparitäts-Element  $EST_i$  jeweils Bilddaten  $\Delta X_{Li}$ ,  $\Delta X_{Rn+1-i}$  ( $1 \leq i \leq n$ ) zugeführt. Diese werden dann einer einzelelementweise (pixelweise) Addition durchführenden Einheit (Bezugszeichen "+") zugeführt, dann einer mit dem Bezugszeichen "1/2" bezeichneten Multiplikationseinheit zugeführt, die eine Multiplikation der addierten Pixelwerte mit einem Faktor 1/2 durchführt, so daß hinsichtlich des jeweils zugeführten Bilddatenpaares eine Mittelwertbildung erfolgt. Daraufaufgehend durchlaufen die derart verarbeiteten Bilddaten ein Ableitungsfilter DIFF.

Ein einfaches Ableitungsfilter DIFF kann beispielsweise Filterkoeffizienten haben, wie sie in der nachfolgenden Tabelle 1 angegeben sind.

Pixelort	x-1	x	x+1
Filterwert	K1 = +0.5	K2 = 0.0	K3 = -0.5

Tabelle 1: Filterkoeffizienten Ableitungsfilter DIFF

Die von diesem Filter ausgegebenen Datenwerte werden einerseits in einer mit dem Bezugszeichen "x<sup>2</sup>" bezeichneten Einheit einer pixelweisen Quadrierung unterzogen, wobei die derart erhaltenen Werte dann einem Mittelungsfiler AVE zugeführt werden.

Andererseits werden die von dem Filter DIFF ausgegebenen Daten einer Multiplikationseinheit (Bezugszeichen "\*") zugeführt, wo sie pixelweise mit der mittels einer Subtraktionseinheit (Bezugszeichen "-") bildelementweise erhaltenen Differenz der Daten  $\Delta X_{Li}$ ,  $\Delta X_{Rn+1-i}$  multipliziert werden. Die derart erhaltenen Werte, die am Ausgang der Multiplikationseinheit zur Verfügung stehen, werden gleichfalls einem weiteren Mittelungsfiler AVE zugeführt.

Für beide Mittelungsfiler AVE können Filterkoeffizienten gewählt werden, wie sie z. B. in der folgenden Tabelle 2 wiedergegeben sind.

Pixelort	x-2	x-1	x	x+1	x+2
Filterwert	J1 = +0.0625	J2 = +0.25	J3 = +0.375	J4 = +0.25	J5 = +0.0625

Tabelle 2: Filterkoeffizienten Mittelungsfiler AVE

Die an den Ausgängen beider Mittelungsfiler erhaltenen Datenwerte werden schließlich einer Divisionseinheit (Bezugszeichen "+") zugeführt, die an ihrem Ausgang den jeweils ermittelten Disparitätswert als Ausgangssignal des jeweiligen Disparitätselements  $EST_i$  ausgibt.

Da für beide eingesetzten Filter lediglich Filterungen in x-Richtung erforderlich sind, was einer Filterung im seriellen Bilddatenverlauf entspricht, kann ein derartiges Disparitäts-Element vollständig in analoger Technik realisiert

werden. Eine digitale Realisierung ist jedoch ebenfalls möglich.

Die derart ermittelten Disparitätswerte, die an den Ausgängen der Disparitätselemente  $EST_i$  ausgegeben werden, werden dann dem Disparitäts-Kohärenzdetektor KD zugeführt.

Fig. 3B zeigt ein Blockschaltbild zur Hardwarerealisierung eines vorstehend mit Bezug auf Fig. 3A funktional beschriebenen einzelnen Disparitätselements  $EST_i$ . Ein Disparitätselement EST besteht dabei aus zur Synchronisation der gesamten Schaltung mit einem von dem Taktgenerator CLK abgeleiteten Taktsignal TAKT getakteten Schieberegistern SR1 bis SR4 zur Zwischenspeicherung einzelner Bildelemente, sowie arithmetischen Logikschaltungen (Mittelwertbildner 1:

" $\Sigma/2$ ", Subtrahierer 2: "-", Multiplikationseinrichtungen 4, 5: MUL, Summierer 3, 6, 7: " $\Sigma$ ", Divisionseinrichtung 8: DIV). Dabei ist das in Fig. 3A mit DIFF bezeichnete Filter nunmehr durch das Schieberegister SR1, einen Koeffizientenmultiplikationsteil K ( $K_1, K_2, K_3$ ) und den Summierer 3 realisiert. Weiterhin sind die in Fig. 3A mit AVE bezeichneten Filter nunmehr durch das Schieberegister SR3, einen Koeffizientenmultiplikationsteil JA (Koeffizienten  $J_1, J_2, J_3, J_4, J_5$ ) und den Summierer 6, bzw. durch das Schieberegister SR4, einen Koeffizientenmultiplikationsteil JB (Koeffizienten  $J_1, J_2, J_3, J_4, J_5$ ) und den Summierer 7 realisiert. Die Disparität wird am Ausgang der Divisionseinrichtung 8 ausgegeben. Die im Blockschaltbild wiedergegebene Schaltungsanordnung verwirklicht dabei vollständig die in Verbindung mit dem Funktionsschaltbild gemäß Fig. 3A angegebene und beschriebene Arbeits- und Wirkungsweise eines Disparitätselements EST.

Fig. 4A zeigt ein Blockschaltbild für einen Kohärenzdetektor bzw. Disparitäts-Kohärenzdetektor KD, wobei als Beispiel ein Kohärenzdetektor KD mit nur vier Eingängen dargestellt ist. Für den in Fig. 2 dargestellten Fall müßte die  
5 Schaltung entsprechend auf sieben Eingänge bzw. allgemein auf  $n$  Eingänge erweitert werden. Falls mehrere ( $z$ ) Verschiebeeinheiten mit jeweils zugeordneten Disparitätsdetektoren DD mit je  $n$  EST Elementen vorgesehen sind, wäre die Schaltung auf  $z \cdot n$  Eingänge zu erweitern.

10

Die von den (vier nicht dargestellten) EST Elementen  $EST_1$  zugeführten und an den Eingängen E1 bis E4 anliegenden Datenwerte der ermittelten Disparitäten werden einer Sortiereinrichtung S1 zugeführt, welche die Datenwerte entsprechend der Größe der Werte sortiert an den Ausgängen A bis D ausgibt. Im dargestellten Fall liegt am Ausgang A der kleinste Wert an, am Ausgang D der größte Wert. Derartige  
15 Sortiereinrichtungen können mit aus Vergleichen

(Komparatoren) und Multiplexern bestehenden Logikgattern  
20 realisiert werden und sind hier nicht im einzelnen erläutert. Jeder der Ausgänge A bis D wird einem Multiplexer MUX1 zugeführt, während weiterhin je zwei der Ausgänge einem Subtrahierer (" - ") zugeführt werden, um die Differenzen B-A, C-B und D-C der entsprechenden Werte zu ermitteln. Die Differenzen werden einer weiteren Sortiereinrichtung S2 zugeführt, die entsprechend der ersten Einrichtung S1 die  
25 Werte der Differenzen betragsmäßig sortiert am Ausgang ausgibt. Dabei ist im vorliegenden Fall nur das kleinste Ergebnis (am Ausgang "<<") weiter zu verarbeiten, welches das

30 Wertepaar der Disparitäten darstellt, die einander am "benachbartsten" sind, also die geringste Abweichung voneinander aufweisen. Das Ergebnis der Sortiereinrichtung S2, also der Wert am Ausgang "<<" wird dann einer auf Gleichheit überprüfenden Einrichtung "=" zugeführt, die das  
35 ermittelte Ergebnis mit den dem Sortierer S2 zugeführten Dif-

ferenzen vergleicht, um zu ermitteln, welcher der Eingangswerte das Ergebnis darstellt. Diese Variante stellt dabei die schaltungstechnisch einfachere Variante dar, denn dies wäre auch aus internen Entscheidungszuständen des Sortierers S2 ermittelbar. Das von der auf Gleichheit überprüfenden Einrichtung erhaltene Ergebnis dient wiederum zur Ansteuerung des Multiplexers MUX1, um die beiden zur Differenz gehörigen Ausgangssignale des Sortierers S1 einem Mittelwertbildner (Bezugszeichen " $\Sigma/2$ ") zuzuführen, der die beiden Werte zur Mittelwertbestimmung addiert und dann durch zwei dividiert. Dieser Mittelwert stellt dann das Ergebnis KE dar, also der Kohärenzwert der Vielzahl ermittelter Disparitätswerte.

Eine Alternative zu der vorstehend beschriebenen Schaltung besteht darin, die ermittelten Differenzen mit einem Schwellenwert zu vergleichen, wobei nur die Kandidaten bzw. Werte weiterverarbeitet werden, die unterhalb des bestimmten Schwellenwerts liegen. Ein nachgeschaltetes Morphologieschaltwerk findet dann den größten Cluster benachbarter Kandidaten. Beide Varianten sind möglich und können bezogen auf den jeweiligen Anwendungsfall jeweils die besseren Ergebnisse liefern.

Fig. 4B zeigt eine Variante des zuvor beschriebenen Kohärenzdetektors KD, bei der dem Kohärenzdetektor neben Ausgangssignalen der EST Elemente weitere Daten zugeführt werden. Diese ist insbesondere dann einsetzbar, wenn die Bilddaten zeilenverschachtelt bzw. "interlaced" erhalten werden, und sich die gesamte Bildinformation eines Vollbildes z.B. aus zwei Halbbildern zusammensetzt. Dann werden die Disparitätswerte jeweils für Bilddaten eines Halbbildes ermittelt und der entsprechende Kohärenzwert der Disparitäten für Bildpunkte eines Halbbildes ermittelt. Für ein jeweiliges Halbbild werden die Werte KE für die Bildpunkte des



Halbbildes in einer Speichereinrichtung MEM (mit einer Speicherkapazität für ein Halbbild) zwischengespeichert, und während der nächsten Halbbildperiode ausgelesen und dem Kohärenzdetektor KD an einem zusätzlich vorzusehenden Eingangsanschluß zugeführt. Der interne Aufbau des Kohärenzdetektors KD ist im wesentlichen identisch zu dem in Verbindung mit Fig. 4A beschriebenen, mit dem Unterschied, daß er ein zusätzliches (das rückgekoppelte) Signal verarbeiten können muß, wofür verschiedene Varianten möglich sind. Insbesondere kann der rückgekoppelte Kohärenz-Disparitätswert bereits dem Sortierer S1 zugeführt werden und dort berücksichtigt werden, oder aber erst später bei der Mittelwertbildung Einfluß nehmen. In letzterem Fall wäre er dem Mittelwertbildner (Bezugszeichen " $\Sigma/2$ ") zuzuführen. Die in Fig. 4A dargestellte Schaltung ist dann jeweils entsprechend zu modifizieren.

Auf diese Weise kann dann der "wahre" Disparitätswert (Kohärenz-Disparitätswert) für ein Vollbild unter Berücksichtigung beider Halbbilder erhalten werden.

Fig. 5 veranschaulicht die Funktionsweise des Disparitäts-Kohärenzdetektors KD. In der graphischen Darstellung gemäß Fig. 5 sind jeweils auf der Abszisse die Bildelementpositionen bzw. Pixelnummern aufgetragen, und als Ordinate für jedes Bildelement die dem Kohärenzdetektor KD zugeführten n Disparitätswerte, die von den Disparitätselementen  $EST_1$  bis  $EST_n$  ausgegeben werden. Im vorliegenden Fall sind beispielsweise  $n = 7$  Disparitätselemente  $EST_1$  für einen Disparitätsdetektor DD vorgesehen.

Für jedes Bildelement werden die vorliegenden n Disparitätswerte mittels des Kohärenzdetektors KD in Echtzeit verglichen und derjenige Disparitätswert als der eigentliche Disparitätswert ausgegeben, der annähernd identisch mit

möglichst vielen weiteren der zugeführten Disparitätswerte ist. Genauer heißt das, daß möglichst viele Disparitätswerte in einem Intervall von z. B. etwa 0.2 bis 0.4 um den betreffenden Disparitätswert herum liegen sollen, damit dieser betreffende Disparitätswert der eigentliche Disparitätswert ist. Am einfachsten ist es, wenn für ein bestimmtes Bildelement mehrere Disparitätswerte den gleichen Wert haben. Dieser Fall entspricht in der graphischen Darstellung gemäß Fig. 5 einem Schnittpunkt von zumindest zwei der n Disparitätswerte für ein bestimmtes Bildelement. Die zeilenweise Analyse der n Disparitätswerte für jedes Bildelement einer betreffenden Zeile liefert dann beispielsweise einen Verlauf der eigentlichen Disparität für eine betreffende Zeile des Bildes wie er im rechten Teil der Fig. 5 dargestellt ist, wobei aus der Gesamtheit der analysierten Bildzeilen dann eine Disparitätskarte in Form eines dritten Bildes ermittelt bzw. dargestellt werden kann (oberes Bild rechts in Fig. 5).

Durch die geeignete Auswahl der Anzahl n von Disparitätselementen EST kann die Genauigkeit bzw. Zuverlässigkeit der seitens des Kohärenzdetektors ermittelten wahren Disparität gesteigert werden, da dann mehr Disparitätswerte vorliegen, die bezüglich ihrer Kohärenz ausgewertet werden. Insbesondere berechnet jedes Disparitäts-Element EST die Disparität im gesamten Meßbereich, also vollständig parallel. Mithin entfällt die bei vielen aus dem Stand der Technik bekannten Realisierungen notwendige iterative Annäherung von groben hin zu feinen Meßbereichen, um den tatsächlichen Disparitätswert zu ermitteln.

Die Auflösung der erhaltenen Disparität im Subpixel-Bereich kann bei analoger Realisierung durch die Anzahl und Abstände der vorzusehenden Abgriffe an den Laufzeitgliedern verändert werden. Bei digitaler Realisierung ist eine entspre-

chende Filterung (Ditherung) der Bilddatenwerte in x-Richtung vorzunehmen, um Bilddaten-Zwischenwerte an Subpixel-Positionen anzunähern.

- 5 Zudem kann beruhend auf den seitens der Diparitäts-Elemente EST des Disparitätsdetektors DD erhaltenen Werte noch eine Zusatzinformation hinsichtlich der Zuverlässigkeit der ermittelten Disparität gewonnen werden. Das heißt, betrachtet man die Disparitätswerte an einem Punkt der Zeile (vgl. Fig. 5), dann ist die Nähe der Nachbarn um den exakten Disparitätswert ein Maß für die Zuverlässigkeit der ermittelten Disparität. Über das gesamte Bild kann somit eine Zuverlässigkeitskarte ermittelt werden, die in Echtzeit eine Aussage bezüglich der Wahrscheinlichkeit für die Richtigkeit der ermittelten Disparität liefert.

- Fig. 6 zeigt ein weiteres Beispiel für von Bildaufnehmern CCDL, CCDD aufgenommene Bilder "linkes Bild" (a) und "rechtes Bild" (b) von im Raum angeordneten Objekten. In beiden Bildern befinden sich ein erstes und zweites Objekt vor einem weiteren (dritten) Objekt, welches im Hintergrund der dargestellten räumlichen Szene zu erkennen ist. Aufgrund der Auswertung durch das erfindungsgemäße optische Sensorsystem zur Verarbeitung stereoskopischer Bilder in Echtzeit wird dann eine Disparitätskarte (c) gewonnen, die als Grauwertbild (oder als Farbbild) auf einem Bildschirm darstellbar ist. Gemäß Fig. 6 ist in der Disparitätskarte ein im Vordergrund befindliches Objekt hell dargestellt, ein im mittleren Tiefenbereich befindliches Objekt durch einen der Tiefe entsprechenden hellen bis dunklen Grauwert dargestellt, und das Objekt im Bildhintergrund als das Objekt mit der größten Tiefe ist in der Disparitätskarte annähernd schwarz wiedergegeben. Die im angegebenen Beispiel gemäß Fig. 6 beschriebene Darstellungsform für die Tiefeninformation ist jedoch nicht zwingend vorgeschrieben. Eben-

so können Objekte im Vordergrund dunkel wiedergegeben sein,  
und Objekte im Hintergrund hell wiedergegeben sein.

Darüber hinaus können die Daten der angezeigten Dispari-  
5 tätskarte auch zu Steuerungszwecken eingesetzt werden, in-  
dem man diese Daten beispielsweise zur Steuerung von Robo-  
tern bei Montagevorgängen heranzieht und einen Soll-Ist-  
Vergleich mit vorab definierten 3-D-Fertigungsdaten und den  
aus den Disparitäten ermittelten Tiefeninformationen durch-  
10 führt.

Unbemannte Fahrzeuge können bei Erreichen und/oder Unter-  
schreiten eines vorbestimmten Minimal-Tiefenwertes zur Ver-  
hinderung von Kollisionen durch entsprechende Ansteuerungen  
15 zum Ausweichen veranlaßt werden, wobei dieses Prinzip auch  
bei der Implementierung passiver und/oder aktiver Abstands-  
warn- und Antikollisionssysteme in Kraftfahrzeugen einsetz-  
bar ist. Weitere Einsatzmöglichkeiten für das System zeigen  
sich in der Medizintechnik beispielsweise beim dreidimen-  
20 sionalen Abtasten von Körperkonturen zur Klassifikation von  
Dysplasien. Insbesondere bestehen Einsatzmöglichkeiten im  
kieferchirurgischen Bereich oder bei flexiblen oder starren  
Endoskopen.



Tiedtke-Bühling-Kinne, POB 20 19 18, D-80019 München

**Patentanwälte  
Vertreter beim EPA\***

Dipl.-Ing. H. Tiedtke\*  
Dipl.-Chem. G. Bühling\*  
Dipl.-Ing. R. Kinne\*  
Dipl.-Ing. B. Pellmann\*  
Dipl.-Ing. K. Grams\*  
Dipl.-Biol. Dr. A. Link\*  
Dipl.-Ing. A. Vollnhals\*  
Dipl.-Ing. T. Leson\*  
Dipl.-Ing. H. Trösch\*  
Dipl.-Ing. Dr. G. Chivarov\*  
Dipl.-Ing. M. Grill\*  
Dipl.-Ing. A. Kühn

**Bavariaring 4,  
D-80336 München**

24. November 1997

DE 20652

**Patentansprüche**

1. Vorrichtung zur Bildverarbeitung, mit

Verschiebe-Einrichtungen (VSEL, VSER;  $\Delta X_{L1}$ , ...,  $\Delta X_{Ln}$ ,

25  $\Delta X_{R1}$ , ...,  $\Delta X_{Rn}$ ) zur gegenseitigen Verschiebung von diesen  
jeweils zugeführten Bilddaten eines ersten Bildes und eines  
zu diesem unter einem anderen Aufnahmewinkel aufgenommenen  
zweiten Bildes;

30 einer Disparitäts-Detektionseinrichtung (DD;  $EST_1$ ,  
...,  $EST_n$ ) zur Erfassung von räumlichen Tiefeninformationen  
wiedergebenden Disparitätswerten für durch die Verschiebe-  
Einrichtungen gegeneinander verschobenen Bilddaten des er-  
sten Bildes und des zweiten Bildes; und

35 einer Kohärenz-Detektionseinrichtung (KD) zur Ermitt-  
lung des tatsächlichen Disparitätswertes unter den Dispari-  
tätswerten dadurch, daß in einem vorbestimmten Wertebereich  
um einen betreffenden der Disparitätswerte eine größtmög-  
liche Anzahl der übrigen Disparitätswerte festgestellt wird.

2. Vorrichtung zur Bildverarbeitung nach Anspruch 1, bei der

die Bilddaten eines Vollbildes während zweier aufeinanderfolgender Halbbildperioden zeilenverschachtelt in

5 Halbbildern zugeführt werden,

von der Kohärenz-Detektionseinrichtung (KD) während einer ersten Halbbildperiode ermittelte tatsächliche Disparitätswerte in einer Speichereinrichtung (MEM) zwischengespeichert werden, und

10 diese zwischengespeicherten Disparitätswerte der Kohärenz-Detektionseinrichtung (KD) während der darauffolgenden zweiten Halbbildperiode zusätzlich zu den von der Disparitäts-Detektionseinrichtung (DD) erfaßten Disparitätswerten zugeführt werden, um die tatsächlichen Disparitätswerte für  
15 das Vollbild zu ermitteln.

3. Vorrichtung zur Bildverarbeitung nach Anspruch 1, bei der

20 die zugeführten Bilddaten des ersten Bildes und des zweiten Bildes jeweils einer Vorverarbeitungseinrichtung (VVL, VVR) zugeführt werden; wobei

aufgrund der Vorverarbeitung erhaltene unterschiedliche Bilddatenarten der Bilddaten des ersten Bildes und des zweiten Bildes jeweils Verschiebeeinrichtungen (VSEL, VSER)  
25 und einer Disparitäts-Detektionseinrichtung (DD) zugeführt werden, die jeweils für eine jeweilige Bilddatenart vorgesehen sind.

4. Vorrichtung zur Bildverarbeitung nach einem der Ansprüche 1 bis 3, bei der

die Bilddaten des ersten und zweiten Bildes seriell und zueinander synchronisiert zugeführt werden.

5. Vorrichtung zur Bildverarbeitung nach Anspruch 4, bei  
35 der die Verschiebe-Einrichtungen (VSEL, VSER;  $\Delta X_{L1}$ , ...,

$\Delta X_{Ln}, \Delta X_{R1}, \dots, \Delta X_{Rn}$ ) die Bilddaten bildpunktweise gegeneinander verschieben.

5 6. Vorrichtung zur Bildverarbeitung nach Anspruch 4, bei der die Verschiebe-Einrichtungen ( $VSEL, VSER; \Delta X_{L1}, \dots, \Delta X_{Ln}, \Delta X_{R1}, \dots, \Delta X_{Rn}$ ) die Bilddaten in Bruchteilen von Bildpunkten gegeneinander verschieben.

10 7. Vorrichtung zur Bildverarbeitung nach Anspruch 5 oder 6, bei der die Bilddaten als analoge Bildsignale zugeführt werden, die Verschiebeeinrichtungen als analoge Laufzeitketten ausgeführt sind und die gegenseitige Verschiebung der Bildelemente durch Einstellung der entsprechenden Laufzeit erfolgt.

15 8. Vorrichtung zur Bildverarbeitung nach Anspruch 5, bei der die Bilddaten als digitale Bildsignale zugeführt werden und die Verschiebeeinrichtungen als getaktete Schieberegister ausgeführt sind.

20 9. Vorrichtung zur Bildverarbeitung nach Anspruch 6, bei der die Bilddaten als digitale Bildsignale zugeführt werden und die Verschiebeeinrichtungen als Filter ausgeführt sind, die Bilddatenwerte zwischen benachbarten Bildelementen interpolieren.

25 10. Vorrichtung zur Bildverarbeitung nach einem der Ansprüche 1 bis 9, bei der einer jeweiligen Disparitätserkennungseinrichtung ( $EST_1, \dots, EST_i, \dots, EST_n, 1 \leq i \leq n$ ) Bilddatenpaare  $\Delta X_{Li}, \Delta X_{Rn+1-i}$  ( $1 \leq i \leq n$ ) zugeführt werden und ein Disparitätswert für diese auf einer gradientenbasierten Verarbeitung ermittelt wird.

11. Vorrichtung zur Bildverarbeitung nach einem der Ansprüche 1 bis 10, bei der

5 zudem eine Anzeigevorrichtung vorgesehen ist, um die tatsächlichen Disparitätswerte als die Tiefeninformation darstellende Disparitätskarte auszugeben.

12. Verfahren zur Bildverarbeitung, mit den Schritten:

10 gegenseitiges Verschieben von zugeführten Bilddaten eines ersten Bildes und eines zu diesem unter einem unterschiedlichen Aufnahmewinkel aufgenommenen zweiten Bildes mittels jeweiliger Verschiebe-Einrichtungen (VSEL, VSER;  
15  $\Delta X_{L1}, \dots, \Delta X_{Ln}, \Delta X_{R1}, \dots, \Delta X_{Rn}$ );

Erfassen von räumliche Tiefeninformationen wiedergebenden Disparitätswerten für die gegeneinander verschobenen  
15 Bilddaten des ersten Bildes und des zweiten Bildes mittels einer Disparitäts-Detektionseinrichtung (DD;  $EST_1, \dots, EST_n$ ); und

Ermitteln des tatsächlichen Disparitätswertes unter den Disparitätswerten derart, daß in einem vorbestimmten  
20 Wertebereich um einen betreffenden der Disparitätswerte eine größtmögliche Anzahl der übrigen Disparitätswerte festgestellt wird.

25 13. Optisches Sensorsystem zur Verarbeitung stereoskopischer Bilder in Echtzeit mit einer Vorrichtung zur stereoskopischen Bildverarbeitung nach einem der Ansprüche 1 bis 11, bei dem die Bilddaten des ersten und zweiten Bildes von zwei in einem vorbestimmten Abstand angeordneten Bildaufnehmern (CCDL, CCCR) zugeführt werden.

30

14. Optisches Sensorsystem zur Verarbeitung stereoskopischer Bilder in Echtzeit nach Anspruch 13, bei dem die erhaltenen Tiefeninformationen zur Steuerung eines beweglichen Objekts im Raum eingesetzt werden.

35





Tiedtke-Bühling-Kinne, POB 20 19 18, D-80019 München

## Patentanwälte Vertreter beim EPA\*

Dipl.-Ing. H. Tiedtke\*  
Dipl.-Chem. G. Bühling\*  
Dipl.-Ing. R. Kinne\*  
Dipl.-Ing. B. Pellmann\*  
Dipl.-Ing. K. Grams\*  
Dipl.-Biol. Dr. A. Link\*  
Dipl.-Ing. A. Vollnhals\*  
Dipl.-Ing. T. Leson\*  
Dipl.-Ing. H. Trösch\*  
Dipl.-Ing. Dr. G. Chivarov\*  
Dipl.-Ing. M. Grill\*  
Dipl.-Ing. A. Kühn

Bavariaring 4,  
D-80336 München

24. November 1997

DE 20652

## Zusammenfassung

Die vorliegende Erfindung betrifft eine Vorrichtung zur Bildverarbeitung, mit Verschiebe-Einrichtungen (VSEL, VSER;  $\Delta X_{L1}, \dots, \Delta X_{Ln}, \Delta X_{R1}, \dots, \Delta X_{Rn}$ ) zur gegenseitigen Verschiebung von diesen jeweils zugeführten Bilddaten eines ersten Bildes und eines zu diesem unter einem anderen Aufnahmewinkel aufgenommenen zweiten Bildes; einer Disparitäts-Detektionseinrichtung (DD;  $EST_1, \dots, EST_n$ ) zur Erfassung von räumlichen Tiefeninformationen wiedergebenden Disparitätswerten für durch die Verschiebe-Einrichtungen gegeneinander verschobenen Bilddaten des ersten Bildes und des zweiten Bildes; und einer Kohärenz-Detektionseinrichtung (KD) zur Ermittlung des tatsächlichen Disparitätswertes unter den Disparitätswerten dadurch, daß in einem vorbestimmten Wertebereich um einen betreffenden der Disparitätswerte eine größtmögliche Anzahl der übrigen Disparitätswerte festgestellt wird. Ferner betrifft die Erfindung ein entsprechendes Verfahren zur Bildverarbeitung.

Fig.2

Telefon: 0 89-54 46 90  
Telefax (G3): 0 89-53 26 11  
Telefax (G4): 0 89-53 29 09 50  
postoffice @ tbk-patent.com

Dresdner Bank (München) Kto. 3939 844 (BLZ 700 800 00)  
Deutsche Bank (München) Kto. 288 1060 (BLZ 700 700 10)  
Postbank (München) Kto. 670-43-804 (BLZ 700 100 80)  
Dai-ichi-Kangyo Bank (München) Kto. 51 042 (BLZ 700 207 00)  
Sanwa Bank (Düsseldorf) Kto. 500 047 (BLZ 301 307 00)

Fig. 1

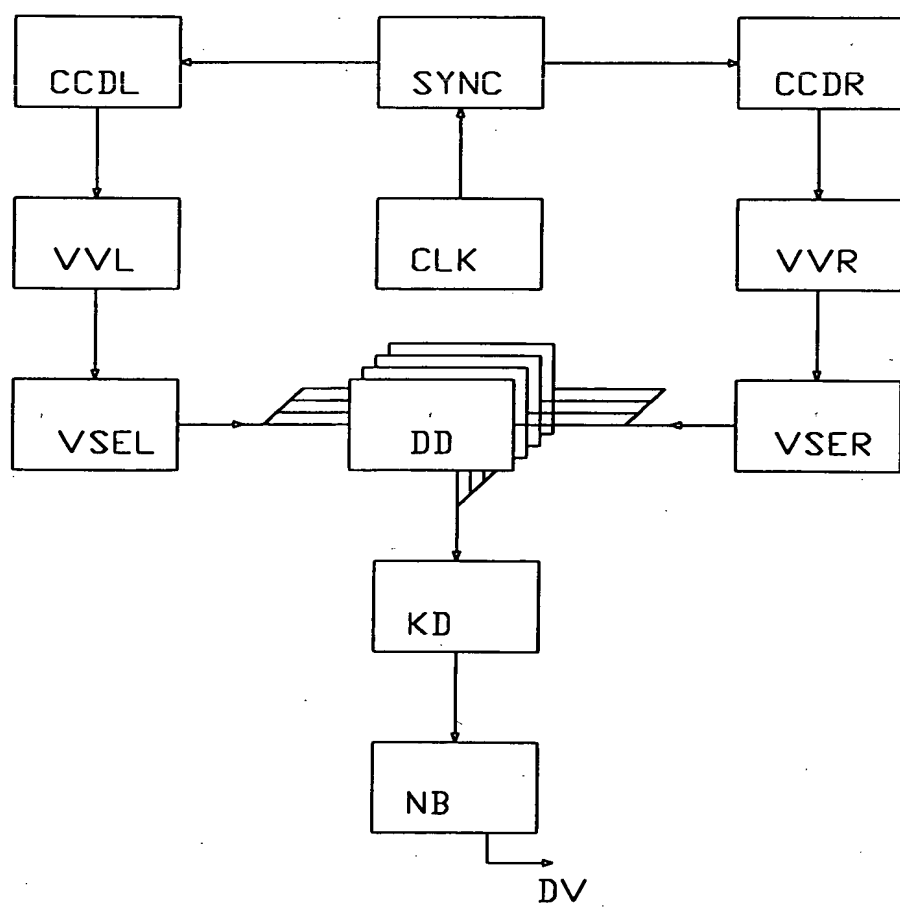


Fig. 2

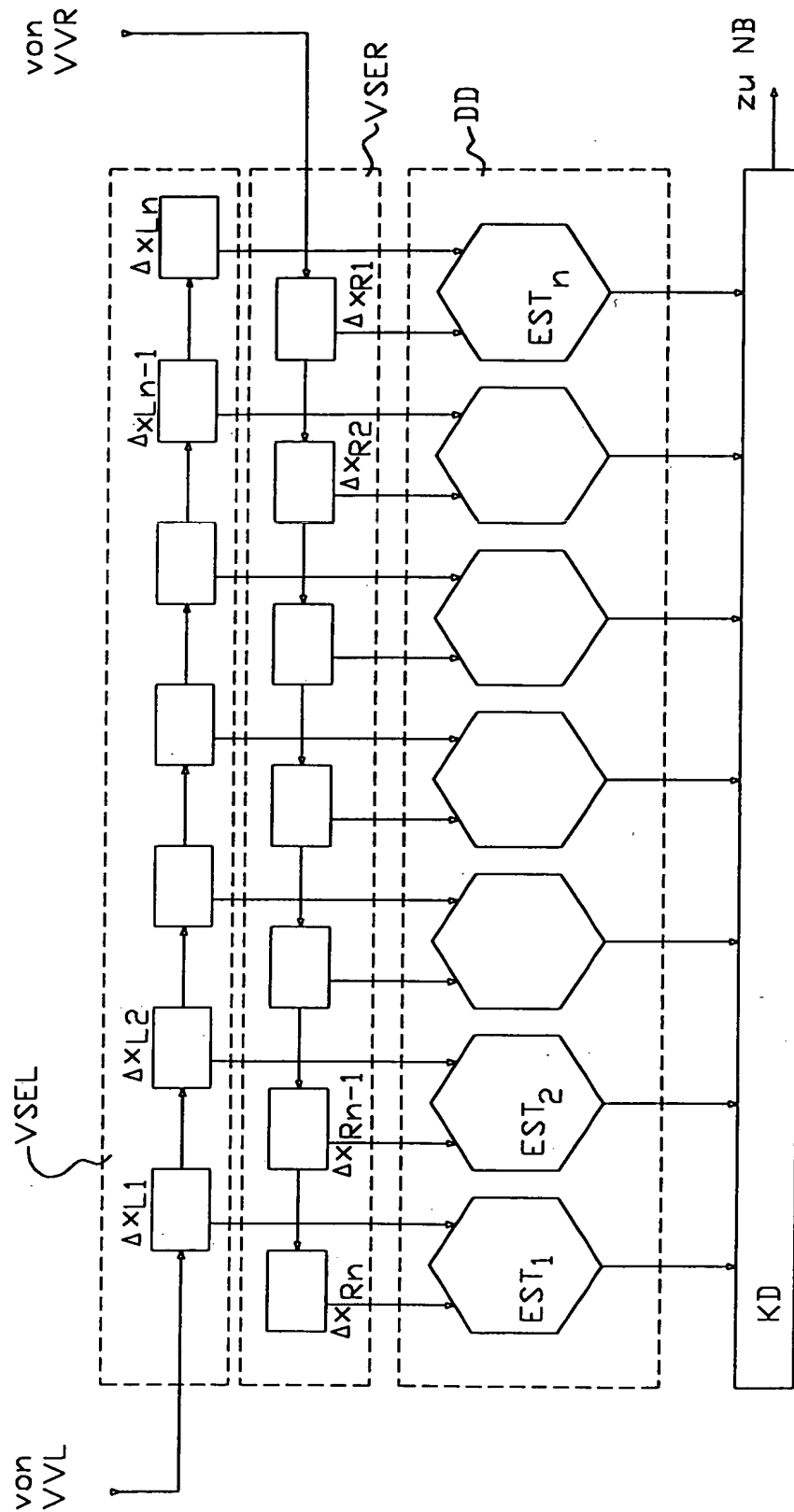


Fig. 3A

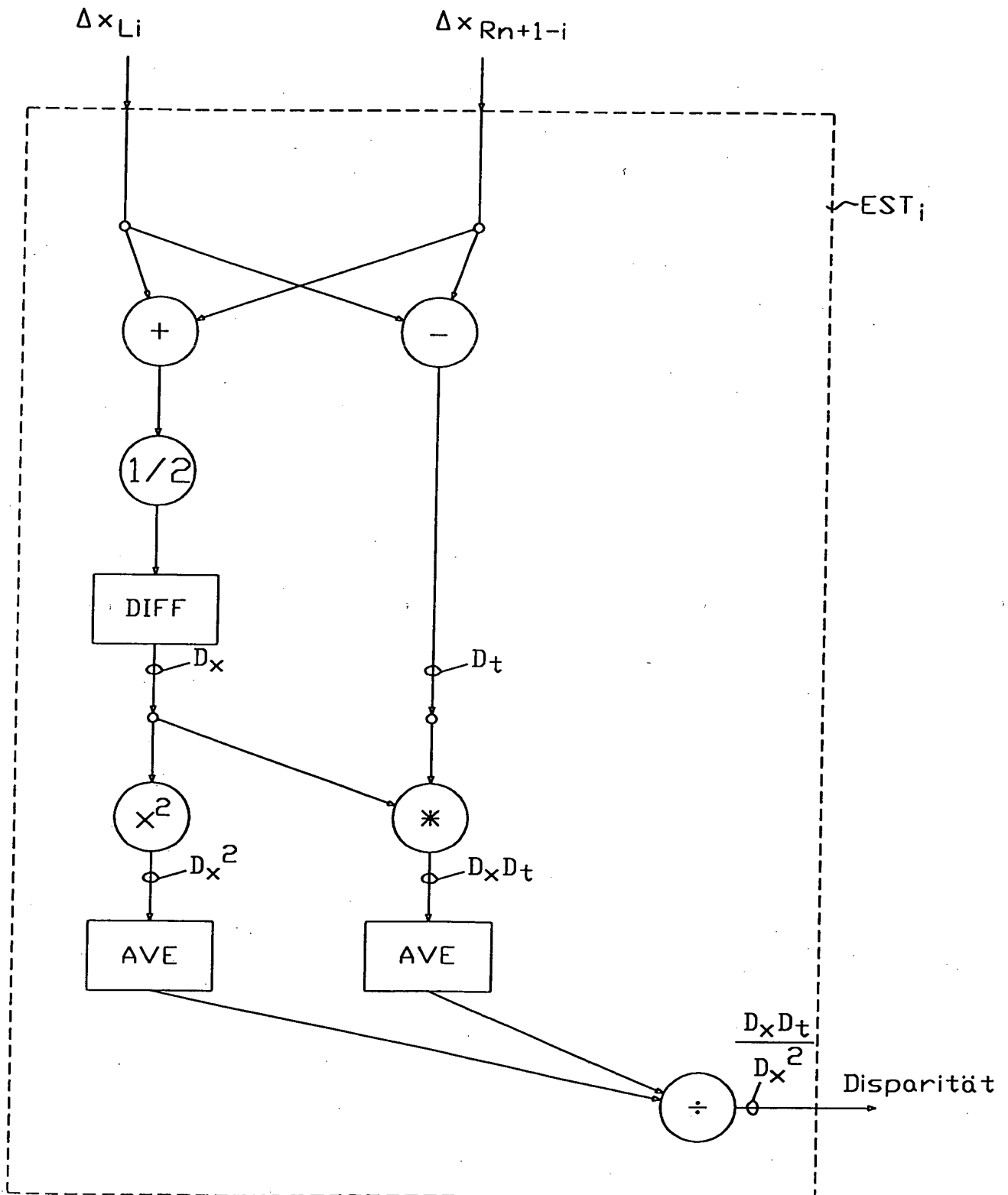


Fig. 3B

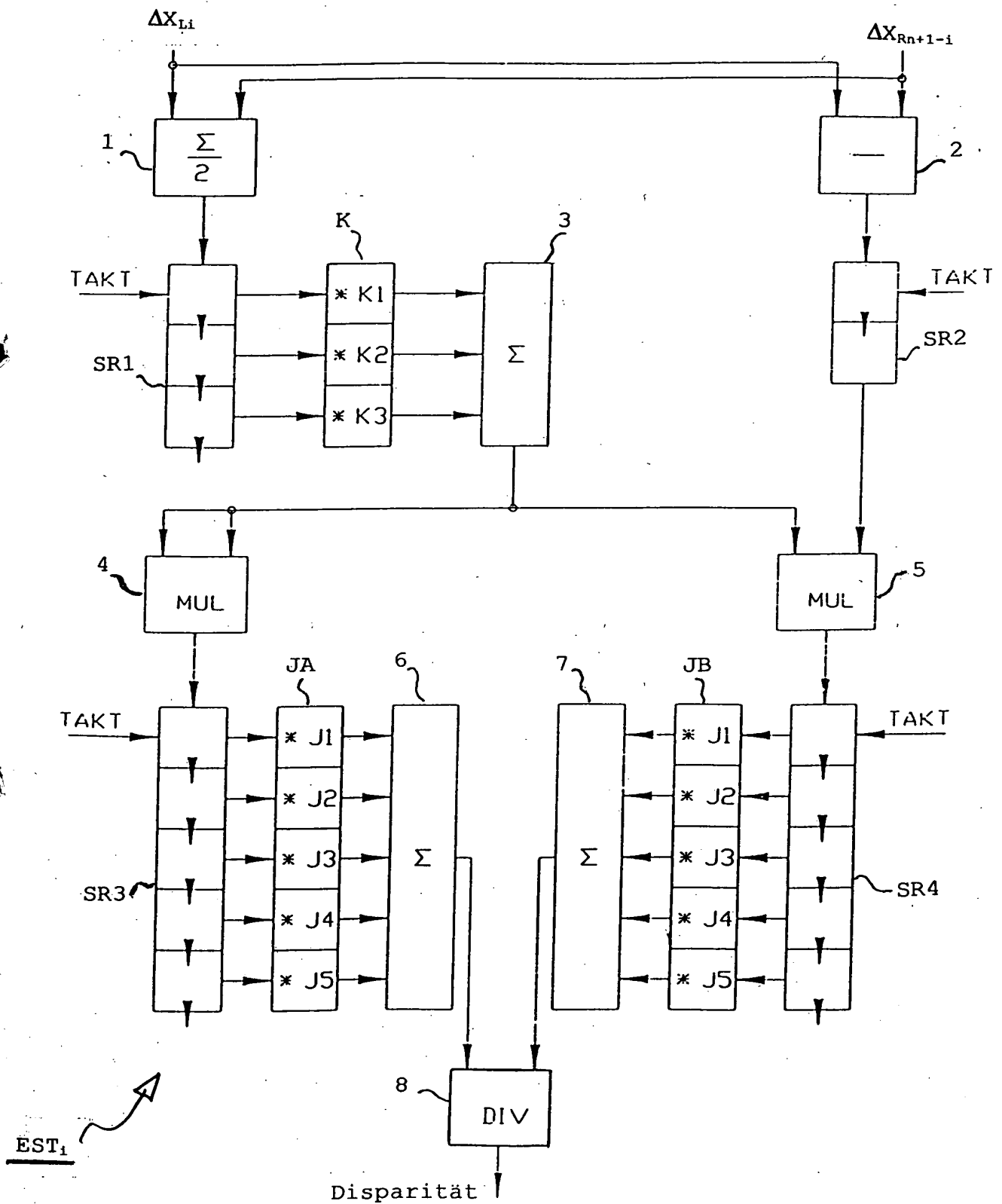


Fig. 4A

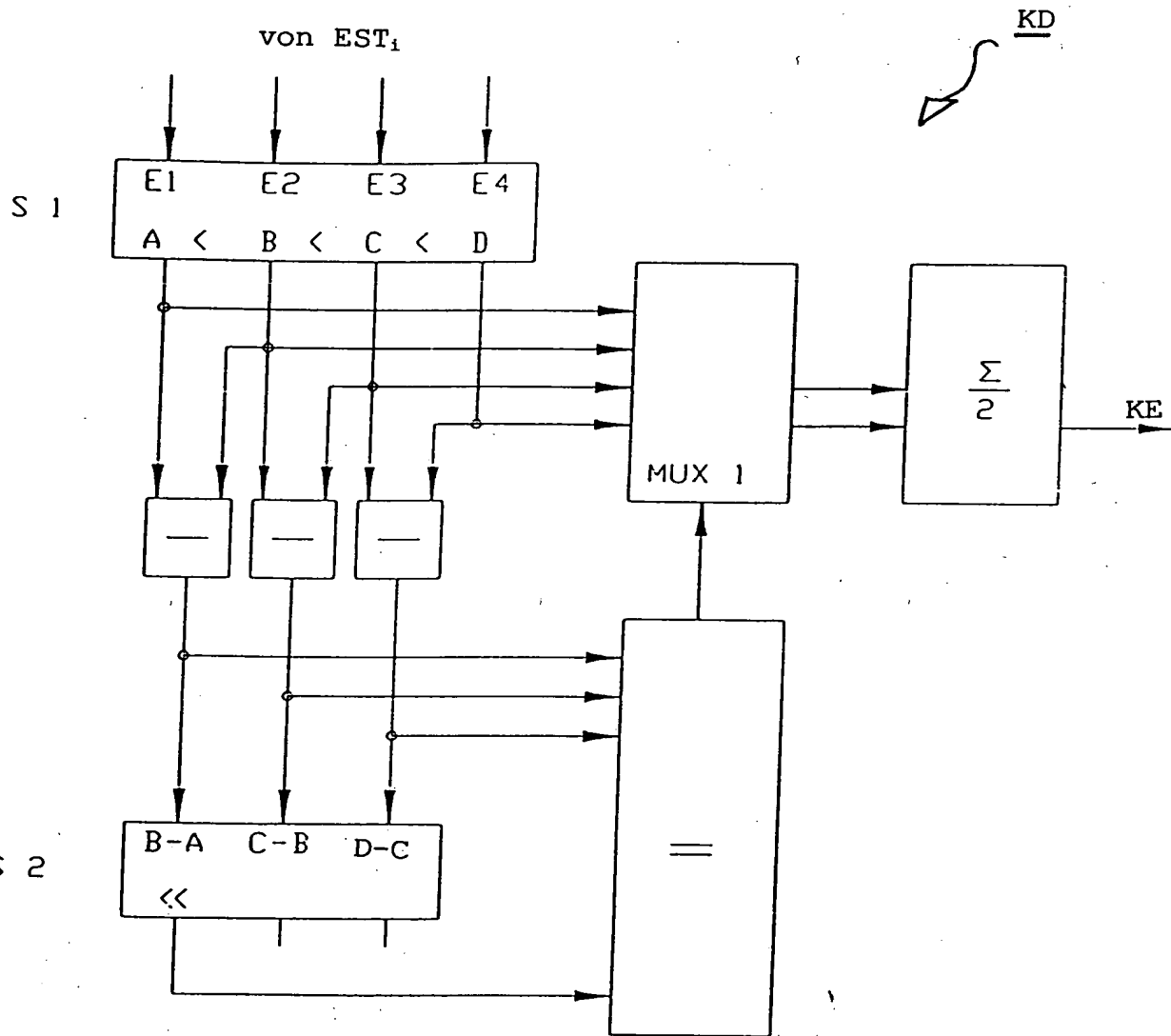


Fig. 4B

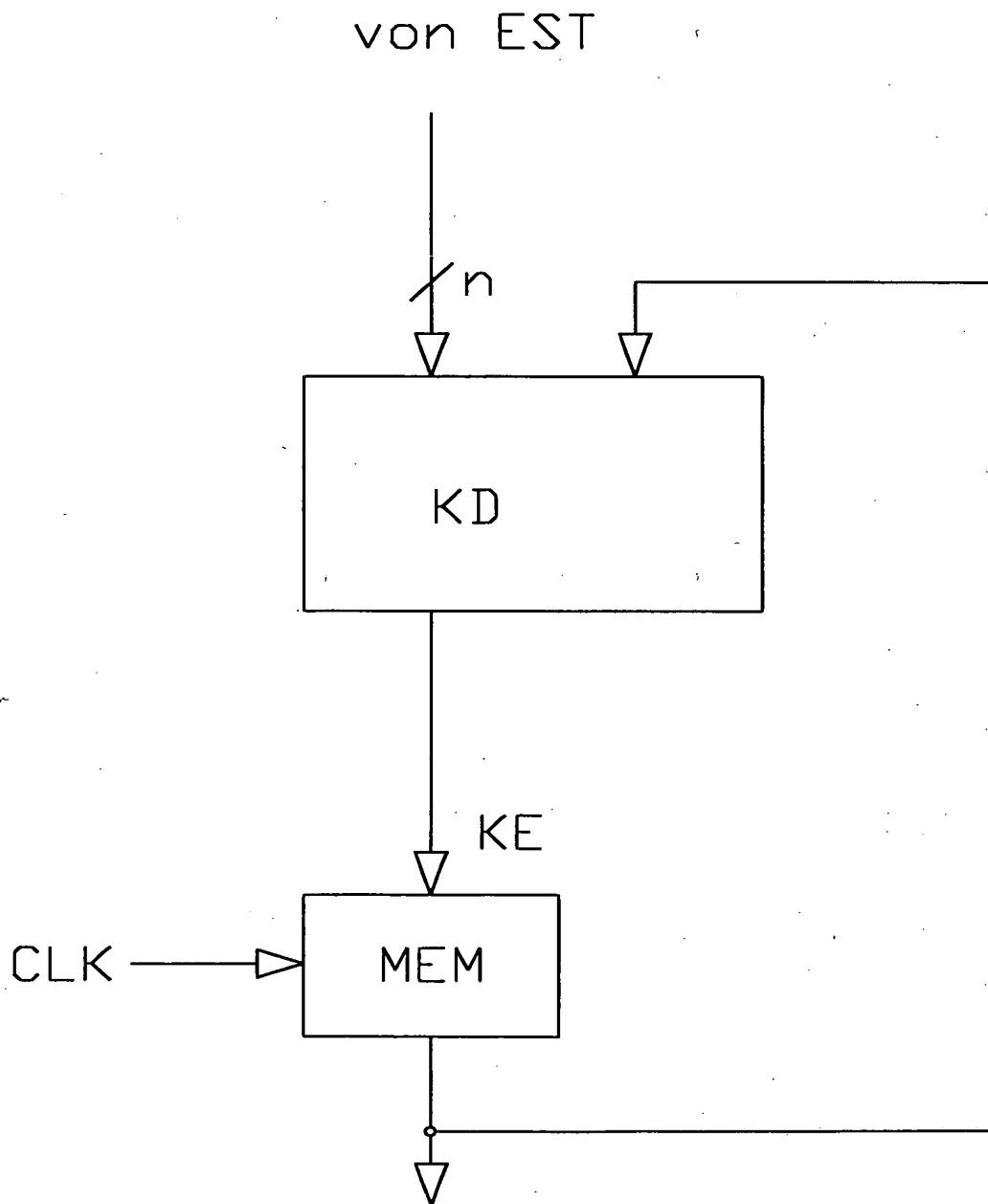


Fig. 5

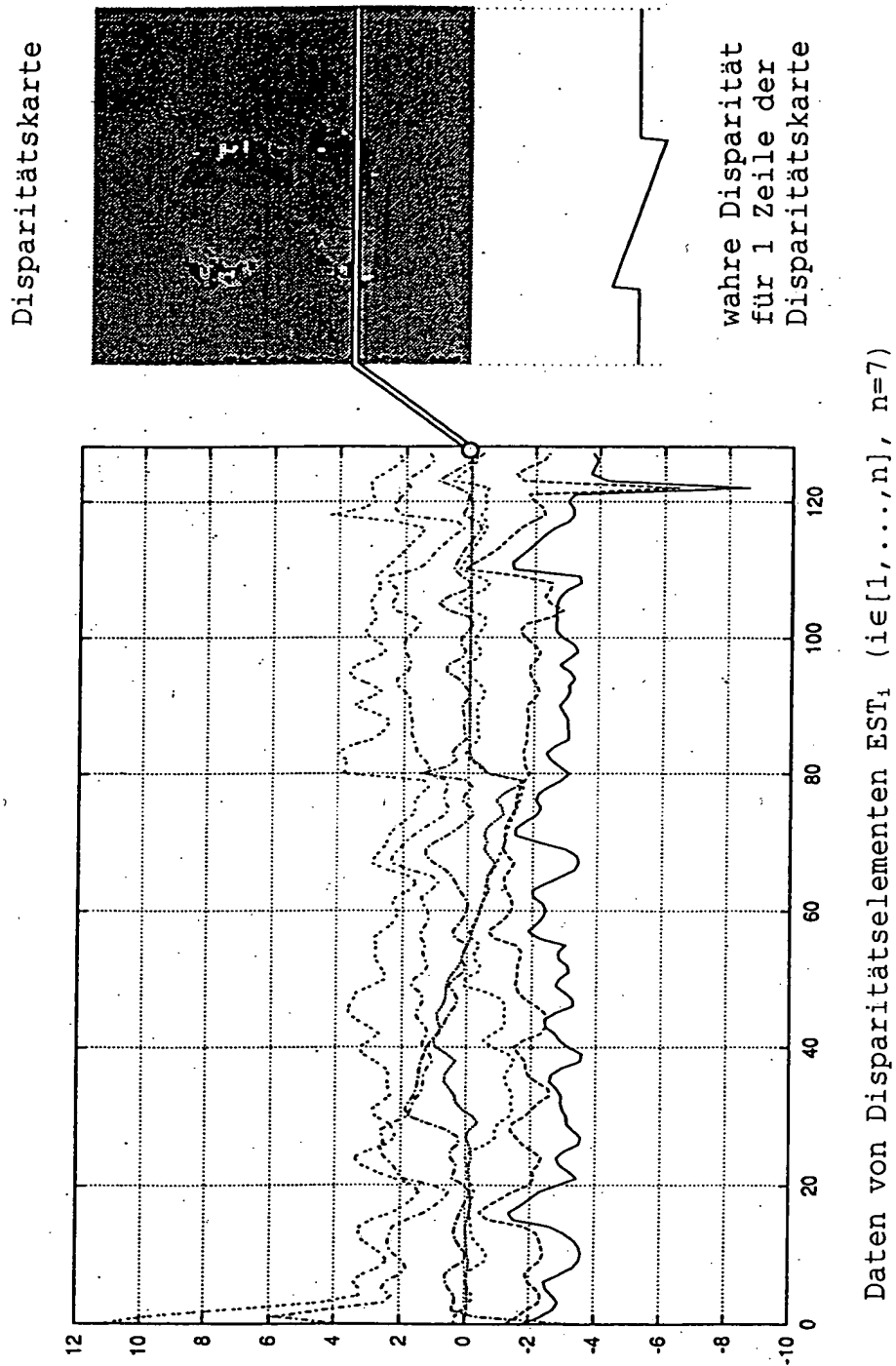
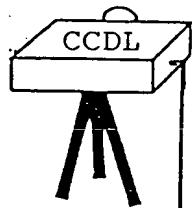
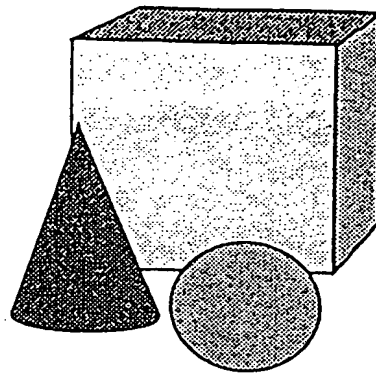


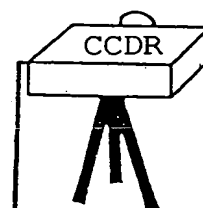
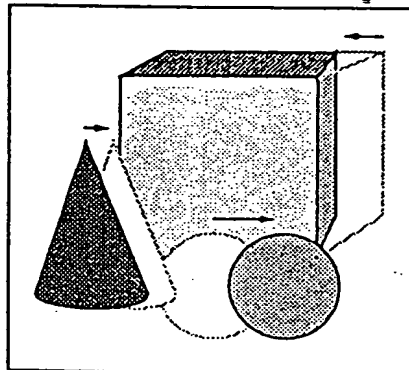


Fig. 6

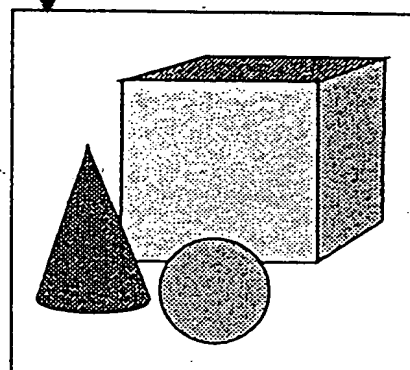
Objekt



linkes Bild (a)



rechtes Bild (b)



Disparitätskarte (c)

